CHAPTER 4

Proposed Fuzzy Logic Traffic Controller for ATM Network

4.1 Introduction

The previous chapter has discussed some of the generic functions for traffic and congestion control in ATM networks. Numerous works has been done to gauge and evaluate the performance of these functions in controlling congestion [22, 23]. As mentioned before, ATM traffic and congestion control functions need to tackle the problem of unpredictable resource demand and the burstiness of data traffic. This couples with different service categories and QoS levels offered by the ATM network poses new challenges not present in packet switched or frame relay network. Most of the time, the traffic control functions such as CAC and UPC have to make decision without sufficient information about the input traffic and the state of the network. Work is still being done in this area to improve the performance of the traffic and congestion control functions.

Recently, fuzzy logic systems have been widely used to control ill-defined and nonlinear systems or processes. Fuzzy logic controllers are known to be extremely flexible and adaptable to changing conditions while reducing complexity in the design. The ability to incorporate expert's knowledge in the fuzzy controllers has made fuzzy logic based controllers very attractive. These characteristics of fuzzy logic controllers are just what ATM researchers are looking for. The previous chapter has presented some of the works that have been done in this area. Currently, there are many papers published on how to utilize fuzzy logic controller in controlling traffic and congestion in ATM networks [3, 5-9].

4.2 Proposed Fuzzy Logic based Traffic Controller

The proposed Fuzzy Logic based Traffic Controller for ATM network consists of two parts:

a) Fuzzy Policer (FP)
b) Fuzzy Congestion Controller (FCC)

A model of the Fuzzy Logic Traffic Controller is presented in Figure 4.1 below:

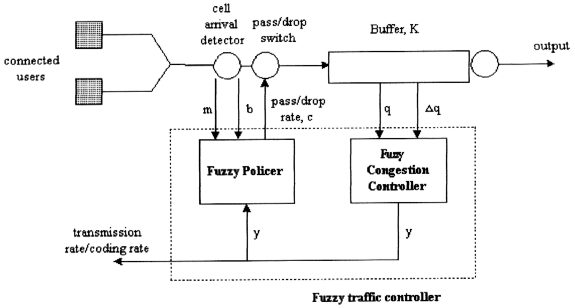

Figure 4.1 Model of Fuzzy Logic Traffic controller

The Fuzzy Policer (FP) is proposed as an alternative to the leaky bucket policing scheme. The FP goal is to simultaneously police mean rate and rejects bursts. Its function is similar to the generic function performed by Usage Parameter Control (UPC). Cells are passed or dropped based on an evaluation of conformance to the traffic contract. Cells are passed if they conform to the traffic contract and negotiated traffic parameters. Cells that are nonconforming will be discarded, or tagged if supported by the network.

There are many reasons for proposing a fuzzy alternative to perform UPC functions. Firstly, current UPC functions are majority performed by the leaky bucket algorithm. The performance of the leaky bucket algorithm has been extensively studied in [22–24]. Previous works have shown that the leaky bucket algorithm has some disadvantages with respect to mean cell rate policing [3] and mean burst length policing [22]. This has encouraged the search for a better algorithm and method to overcome these inefficiencies.

Secondly, considerable works have been done on utilizing fuzzy based approach in other areas of traffic and congestion control functions. For example, applying fuzzy systems to estimate effective bandwidth for connection admission control functions has been done in [25]. Another example is the use of fuzzy system to estimate cell loss ratio for applications in connection admission control functions [8]. Both have produced promising and good results. Fuzzy based rate control scheme has also been proposed for the delivery of real-time MPEG video in ATM networks [26]. Therefore, it is best to work on UPC function to emulate the success of fuzzy controllers in these areas. In addition, intense research is still being done on developing new algorithms and approach for UPC functions.

The Fuzzy Congestion Controller (FCC) is a Fuzzy Logic Controller (FLC) that can prevent or relieve network congestion. The FCC is a fuzzy implementation of the two threshold congestion control methods [27-29]. Current implementation of the two threshold congestion control methods has the difficulty of determining the effective thresholds under various bursty traffic conditions in ATM networks. Therefore, a fuzzy based approach is proposed as a solution to control congestion.

4.2.1 Fuzzy Policer (FP)

A fuzzy logic based controller is proposed to perform usage parameter control (UPC) functions. This fuzzy controller is called the Fuzzy Policer (FP) and its main goal is to police the mean rate while at the same time rejects bursts.

The proposed FP is an enhancement over the fuzzy policer suggested in [7]. The FP in this thesis is an extension over the original fuzzy policer. The original fuzzy policer is a two-input, nine-rule, one-output system. In contrast, the FP presented in this thesis is a three-input, eighteen-rule, one-output system. The reason for the extension is to provide the FP with more information thus helping it in making more accurate decisions on passing /dropping cells.

The objective of the FP is to police the mean rate and reject bursts. The FP performs its job by continuously evaluating the compliance/violation level of two parameters:

a) ratio of up-to-date mean bit rate to negotiated mean bit rate **(1)**
b) ratio of up-to-date mean burst length to negotiated mean burst length

(1) mean cell rate can also be used

It then decides on the drop rate to be used on the cells based on the collective evaluation of the compliance/violation level of the two parameters. Policing of mean rate is simple, cells are discarded or tagged when the mean rate is above the negotiated mean rate. However, dealing with violation of burst length requires more information. When policing the burst length, the FP needs to acquire the mean rate together with the burst type before taking actions on discarding cells. The FP has to evaluate both the violation level of mean rate and burst length to decide on a proper action.

For example, consider the case where the mean rate is below the negotiated mean rate, but the burst length is slightly above the negotiated burst length. If the decision to discard cells is made based on violation of burst length, then of course it would improve the burstiness of cell stream. However, discarding cells will also reduce the throughput capacities of the mean rate. Therefore, violation of burst length in this situation may need to be excused. On the other hand, if the same situation arises but the burst length is much greater than the negotiated value, then some penalty should be imposed by discarding some fraction of the cells. If both mean rate and burst length are above the negotiated values, a greater punishment is given by discarding more cells. In short, the FP will decide on the action to be taken by evaluating the compliance/violation level of both mean rate and burst length.

The proposed FP in this thesis adds another parameter as input to provide more information in making decisions. The parameter comes in the form of an indicator of the state of the network. This parameter will provide indication to the FP on the congestion state of the network. It informs the FP on whether the network is in a congested state or not. If the network is expecting congestion or is congested, then discarding cells is more favorable to reduce the intensity and prevent the spread of congestion. Looking back at the situation above, where the mean rate is below negotiated rate and burst length slightly above negotiated values, passing cells may minimize the reduction of throughput. However, if the network is congested, then passing cells will only lead to more cells being dropped due to buffer overflow and worsen the congestion. In the end, the throughput will drop as well. Thus, discarding some fraction of cells during congestion period is more appropriate.

Hence, the proposed FP has the advantage of acquiring more information about the input traffic and the state of the network before making any decisions. The proposed FP will continuously evaluate the compliance/violation level of both mean rate and burst length, and the state of network to decide on passing or dropping cells. If congestion occurs, the FP will impose a heavier penalty than during congestion free, depending on the violation level of the parameters. Equipped with these information, the proposed FP is able to reduce uncertainties and thus make more accurate decisions in policing cells.

As for the leaky bucket algorithm, previous works have shown that it has problems in policing mean cell rate [3] and mean burst length [22]. The reason for less than satisfactory results in policing mean cell rate is that the characteristics of the source must be estimated on the basis of relatively short sample. This causes a certain probability of incorrect policing decisions.

In addition, mean rate policing requires setting the leak rate equals to or slightly greater than negotiated mean rate while the bucket size should be large to accommodate short bursts. In this setting, the leaky bucket algorithm is sensitive to mean rate violation. However, a large bucket size is not desirable because of long response time and slow reaction to long bursts. Another way to police mean rate is to set the leak rate greater than negotiated mean rate with a smaller bucket size. A smaller bucket size improves on detecting long bursts, but a larger leak rate compared to negotiated mean rate reduces ability of the algorithm to efficiently police mean rate and short bursts. Therefore, the leaky bucket algorithm is ineffective in policing bursty traffic.

The proposed FP offers a solution to the first problem by utilizing up-to-date value of the mean rate and burst length. The up-to-date value is calculated by taking into account all the values for each parameter (mean rate and burst length) respectively thus extending the sampling size. This enables more accurate information and reduces probability of incorrect policing decisions.

As for bursty traffic, the proposed FP collectively evaluates both the compliance/violation level of mean rate and burst length before making policing decisions. Therefore, the FP is able to make a better decision to either pass or drop cells in different bursty situation. Furthermore, the FP also considers the state of the network which is an important piece of information when making policing decisions.

Take note that the leaky bucket algorithm requires completely different dimensioning for peak rate and mean rate policing [22, 23]. To police mean rate, the leak rate should be set slightly lower than negotiated mean rate while the bucket size needs to be large, so that decision is made based on sufficiently long samples of incoming cells. However, this setting is unacceptable for peak rate policing because a burst will easily exceed the leak rate even though it is within negotiated values. To police peak rate, the leak rate should be set near the negotiated peak rate while the bucket size needs to be small, so that decision can be made rapidly while the burst is still on. However, violation of mean rate will go undetected because the leak rate is much higher than the violated actual mean rate.

The next section will look into details of the FP, including its inputs, rule base and output control action.

The proposed FP is a three-input, eighteen-rule, one-output system. The three-input values to FP are:

a) ratio of up-to-date mean bit rate to negotiated mean bit rate (A_1) **(1)**

b) ratio of up-to-date mean burst length to negotiated burst length (A_2)

c) state of the network (y)

(1) mean cell rate can also be used

The up-to-date value is calculated by averaging all previous values up to the most current value. For example, if we have m_1, m_2, m_3, ...m_i as the values of actual mean rate, then

$$\text{up-to-date mean rate} = \frac{1}{i}\sum_{1}^{i} m_i$$

The ratio of up-to-date mean bit rate to negotiated mean bit rate and the ratio of up-to-date mean burst length to negotiated burst length serves as indication of the compliance /violation level of the two parameters. For example, a higher value of the ratio shows a greater violation of the up-to-date value over the negotiated value. A lower ratio, for example, having a value of less than or equal to one, shows full compliance with the negotiated value.

The value of input y, which represents the state of the network, may be positive or negative. If y is a negative value, then it indicates that there is congestion. On the other hand, a positive value of y implies congestion free environment.

The output value of the FP is the drop rate c, that is sent to a pass/drop switch which will then pass or discard a portion of the cells held in the memory based on the value of c. If the FP detects violation of the negotiated rate or congestion, then the switch will discard cells. The output value of the FP is:

a) drop rate (c)

These three inputs are first fuzzified, then fed to the rule base in the inference engine, and later defuzzified to produce a crisp output action c, which dictates the drop rate to be performed on the cells. The output value c, is within the range of [0, 1] whereby 0 denotes total drop of all cells and 1 denotes total pass of all cells.

The FP will evaluate the compliance/violation levels of the two inputs and the state of the network, which is provided by the third input to calculate the drop rate. The third input is generated by the FCC. The drop rate value is then sent to a pass/drop switch. The switch should have enough memory to store incoming cells before a decision is made. The switch will then either pass or drop a fraction of the cells in the memory based on the values of the drop rate.

The following notations are used:

A_1 = ratio of up-to-date mean bit rate to negotiated mean bit rate

A_2 = ratio of up-to-date mean burst length to negotiated burst length

y = state of the network

c = drop rate

The term set used to describe each input parameter and the output drop rate is defined as follows:

Term set for A_1, $T(A_1)$ = { Violate (V), Sort of Comply (SC), Comply (C) }

Term set for A_2, $T(A_2)$ = { Violate (V), Sort of Comply (SC), Comply (C) }

Term set for y, T(y) = { Negative (N), Positive (P) }

(negative - congestion, positive - no congestion)

Term set for drop rate c, T(c) = {Drop (D), Between Pass & Drop (B), Pass (P)}

The membership functions for the terms in the term set should be defined with the proper shape and position. Usually, the shapes of the membership function used are triangular or trapezoidal functions because these functions are suitable for real-time operation [14].

The triangular function and trapezoidal function are defined as follows:

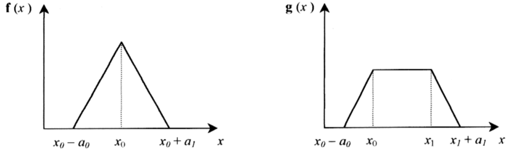

Figure 4.2 Triangular and Trapezoidal membership functions

$$f(x : x_0, a_0, a_1) = \begin{cases} \dfrac{x - x_0}{a_0} + 1 & \text{for } x_0 - a_0 < x \le x_0 \\ \dfrac{x_0 - x}{a_1} + 1 & \text{for } x_0 < x \le x_0 + a_1 \\ 0 & \text{otherwise} \end{cases}$$

$$g(x : x_0, x_1, a_0, a_1) = \begin{cases} \dfrac{x - x_0}{a_0} + 1 & \text{for } x_0 - a_0 < x \le x_0 \\ 1 & \text{for } x_0 < x \le x_1 \\ \dfrac{x_1 - x}{a_1} + 1 & \text{for } x_1 < x \le x_1 + a_1 \\ 0 & \text{otherwise} \end{cases}$$

The membership functions for the term set $T(A_1)$ and $T(A_2)$ are illustrated in Figure 4.3(a) and (b) below:

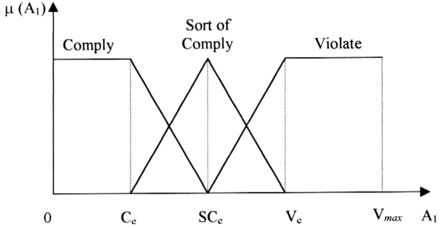

Figure 4.3(a) The membership functions for the term set $T(A_1)$

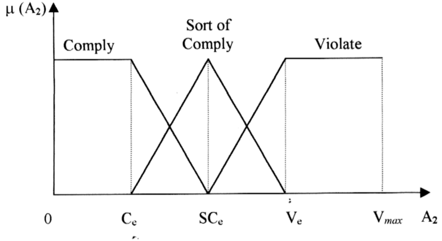

Figure 4.3(b) The membership functions for the term set $T(A_2)$

The membership functions for terms C, SC, V in the term set $T(A_1)$ and $T(A_2)$ are given by:

$$\mu_C(A_1) = g\ (A_1 : 0,\ C_e,\ 0,\ C_w)$$

$$\mu_{SC}(A_1) = f\ (A_1 : SC_e,\ SC_w,\ SC_w)$$

$$\mu_V(A_1) = g\ (A_1 : V_e,\ V_{max},\ V_w,\ 0)$$ *V_{max} = the largest possible violation ratio

$\mu_C(A_2) = g\ (A_2 : 0,\ C_e,\ 0,\ C_w)$

$\mu_{SC}(A_2) = f\ (A_2 : SC_e,\ SC_w,\ SC_w)$

$\mu_V(A_2) = g\ (A_2 : V_e,\ V_{max},\ V_w,\ 0)$ *V_{max} = the largest possible violation ratio

C_e represents the maximum ratio value of the parameter (A_1, A_2) that can still be considered to be compliant. SC_e provides a soft transition value between compliance and violation. V_e is the ratio value that is considered to be in violation.

The membership functions for the term set T(y) are illustrated in Figure 4.4 below:

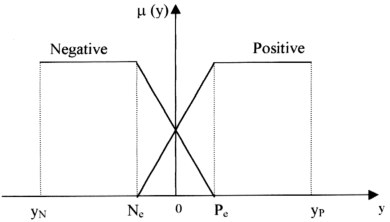

Figure 4.4 The membership functions for the term set T(y)

The membership functions for terms N, P in the term set T(y) are given by:

$\mu_N(y) = g\ (y : y_N,\ N_e,\ 0,\ N_w)$ *for congestion

$\mu_P(y) = g\ (y : P_e,\ y_P,\ P_w,\ 0)$ *for no congestion

y_N represents the maximum negative value that the input y can take, whereas y_P represents the maximum positive value for y. The value between N_e and P_e provides a soft transition between congestion and congestion free state.

The membership functions for the term set T(c) are illustrated in Figure 4.5 below:

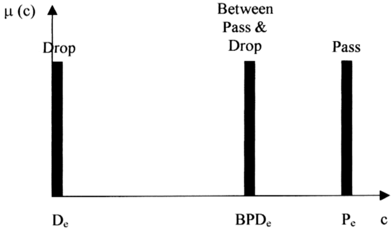

Figure 4.5 The membership functions for the term set T(c)

The membership functions for terms D, BPD, P in the term set T(y) are given by:

$\mu_D(c) = f\ (c : D_e\ ,\ 0\ ,\ 0\ )$

$\mu_{BPD}(c) = f\ (c : BPD_e\ ,\ 0\ ,\ 0\ )$

$\mu_P(c) = f\ (c : P_e\ ,\ 0\ ,\ 0\ )$

D_e, BPD_e, P_e represents the drop rate imposed on the cells. D_e would be set to 0 for total drop of all cells, BPD_e to a value within [0, 1] but closer to 1 for dropping a fraction of the cells, and P_e is set to 1 for passing all cells.

The FP decides on the drop rate c to either pass or drop cells, according to the set of linguistic variables of parameters A_1 and A_2, state of the network y generated by FCC, and a set of built in fuzzy control rules. The rule base is constructed on the basis of knowledge from policing experts.

The rule base consists of a set of linguistics rules in the form of "if – then" statement that describes the fuzzy logic relationship between the input variables and the output variables. The rules are read as follows:

IF A_1 is C **and** A_2 is C **and** y is P **THEN** c is P

The eighteen–rule structure for the Fuzzy Policer is presented in Table 4.1 below:

Rule	A_1	A_2	y	c	Rule	A_1	A_2	y	c
1	C	C	P	P	10	C	C	N	P
2	C	SC	P	B	11	C	SC	N	B
3	C	V	P	D	12	C	V	N	D
4	SC	C	P	B	13	SC	C	N	D
5	SC	SC	P	B	14	SC	SC	N	D
6	SC	V	P	D	15	SC	V	N	D
7	V	C	P	D	16	V	C	N	D
8	V	SC	P	D	17	V	SC	N	D
9	V	V	P	D	18	V	V	N	D

Table 4.1 Rule Structure for Fuzzy Policer (FP)

The proposed FP uses the max-min inference method [14] discussed in section 2.4.3 for the inference engine because it is designed for real-time operation. As for the defuzzification process, the FP uses the Tsukamoto's Defuzzification Method (discussed in section 2.4.4) because of its simplicity in computation. The defuzzification process produces a crisp value of the drop rate, c that will be used to calculate the number of cells to pass/drop.

4.2.2 Fuzzy Congestion Controller (FCC)

The Fuzzy Congestion Controller (FCC) is a fuzzy implementation of the two threshold congestion control method studied in [27-29]. One of the difficulties of current two threshold congestion control method is determining the effective thresholds under various bursty traffic conditions in ATM networks. Therefore, a fuzzy logic approach is proposed to tackle this problem.

The proposed FCC is a two-input, four-rule, one-output system. The two-input values to FCC are:

a) queue length of buffer (q)

b) queue-length change rate (Δq)

The FCC will continuously monitor the queue length of the buffer to check whether the queue length has exceeded the threshold values or not. Two threshold values are determined, one for the high threshold and the other for the low threshold. These threshold values are used to determine the onset and the relief of congestion. If the queue length exceeds the high threshold value, it means that the system is experiencing congestion. On the other hand, if the queue length drops below the low threshold, it means that the system is congestion free or has just recovered from congestion. The FCC also takes into account the queue-length change rate as an indicator of the state of the system. This indicator provides extra and more accurate information to the FCC in determining the congestion state of the system. By evaluating the buffer queue length and the queue-length change rate, the FCC offers a more accurate view of the state of the system.

The output value of the FCC is the rate control y, which is sent to the source. This rate control y, acts as a signal to notify the source to adjust its current transmission (cell) rates based on the congestion state of the system. If the system is experiencing congestion, the FCC will sent a negative value of y to inform the source to decrease its current rate. The source can then decrease its current rates by reducing transmission rates or selectively discarding cells. If the system is congestion free, then a positive value of y is sent, so that the source can restore to its original rate. This step is necessary to improve throughput.

The output value of the FCC is:

a) rate control (y)

The following notations are used for the FCC:

q = queue length of buffer

Δq = queue-length change rate

y = rate control

The term sets used to describe each input parameter and the output control rate are defined as follows:

Term set for queue length T(q) = { Empty (E), Full (F) }
Term set for queue-length change rate T(Δq) = { Negative (N), Positive (P) }
Term set for rate control T(y) = { Decrease (DC), No Change (NC), Increase (I) }

The terms Empty and Full are used to describe the queue length. As for the queue-length change rate, the term Negative indicates that the occupancy of the queue is reducing while the term Positive shows that the occupancy is increasing. The terms Decrease, No Change and Increase are used to describe the rate control action generated by the FCC.

The membership functions for the term set T(q) are illustrated in Figure 4.6 below:

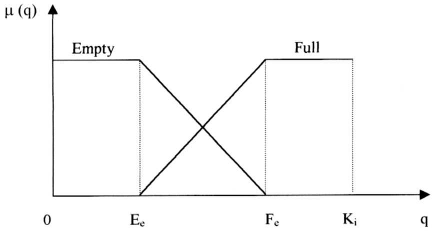

Figure 4.6 The membership functions for the term set T(q)

The membership functions for terms E, F in the term set T(q) are given by:

$$\mu_E(q) = g\ (q : 0,\ E_e,\ 0,\ E_w)$$
$$\mu_F(q) = g\ (q : F_e,\ K_i,\ F_w,\ 0)$$

The values at E_e and F_e represent the low threshold and high threshold values respectively, as used in the two threshold congestion control. A value of q below low threshold indicates a congestion free state while a value of q above the high threshold implies congestion. K_i represents the total buffer size.

The membership functions for the term set T(Δq) are illustrated in Figure 4.7 below:

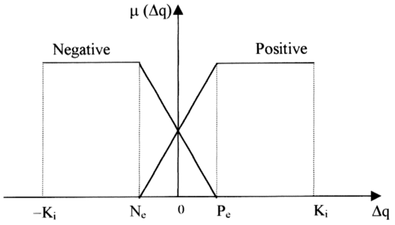

Figure 4.7 The membership functions for the term set T(Δq)

The membership functions for terms N, P in the term set T(Δq) are given by:

$\mu_N(\Delta q) = g\ (\Delta q : -K_i,\ N_e,\ 0,\ N_w)$

$\mu_P(\Delta q) = g\ (\Delta q : P_e,\ K_i,\ P_w,\ 0)$

$-K_i$ represents the maximum possible value that Δq can take during congestion free period, whereas K_i represents the maximum value for Δq during congestion.

The membership functions for the term set T(y) are illustrated in Figure 4.8 below:

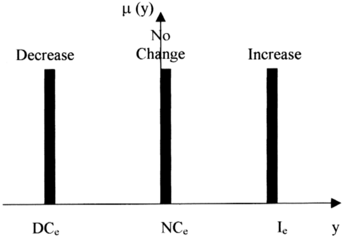

Figure 4.8 The membership functions for the term set T(y)

The membership functions for terms DC, NC, I in the term set T(y) are given by:

$\mu_{DC}(y) = f\ (y : DC_e\ ,\ 0\ ,\ 0\ )$

$\mu_{NC}(y) = f\ (y : NC_e\ ,\ 0\ ,\ 0\ )$

$\mu_{I}(y) = f\ (y : I_e\ ,\ 0\ ,\ 0\ )$

DC_e, NC_e, I_e represent the rate control sent to the source. NC_e would be set to 0, DC_e and I_e are set to a negative and positive value respectively. They would be set to the same magnitude for symmetry.

The FCC generates the rate control action y according to the set of linguistic variables of queue length, q and queue-length change rate, Δq and a set of built in fuzzy control rules. The rule base is constructed on the basis of knowledge of the two-threshold congestion control method.

The rule base consists of a set of linguistics rules in the form of "if – then" statement that describes the fuzzy logic relationship between queue length and queue-length change rate with the output rate control y. The rules are read as follows:

IF q is E **and** Δq is N **THEN** y is I

The rule structure for the FCC is presented in Table 4.2 below:

Rule	q	Δq	Y
1	E	N	I
2	E	P	I
3	F	N	D
4	F	P	D

Table 4.2 Rule Structure for Fuzzy Congestion Controller (FCC)

The two inputs are first fuzzified, then fed to the rule base in the inference engine, and later defuzzified to produce a crisp output. The proposed FCC also uses the max-min inference method [14] discussed in section 2.4.3 for the inference engine. The Tsukamoto's Defuzzification Method (discussed in section 2.4.4) is utilized for the defuzzification process. Finally, a crisp rate control value is produced and sent to the source. The source then adjust its transmission (cell) rate accordingly to prevent or to relieve congestion.

4.3 Operation of the Fuzzy Logic Traffic Controller

Input traffic from connected users is fed into the switch before being admitted to the buffer. A cell arrival detector measures the input mean bit rate (m) and input mean burst length (b). Incoming cells are stored temporarily in the pass/drop switch. The two measurement m and b serve as inputs to a Fuzzy Policer, which decides whether the incoming cells comply with the traffic contract. The Fuzzy Policer also takes into account the condition of the network (whether congested or not) provided by the Fuzzy Congestion Controller in making the decision. The output decision is then forwarded to the pass/drop switch. It is the job of the pass/drop switch to either pass cells that comply or drop cells that violate the traffic contract. Cells that are passed are then admitted to the buffer

The Fuzzy Congestion Controller accepts the queue length (q), and queue-length change rate (Δq) of the buffer as inputs and output a rate control action (y) to traffic sources.

The source can then increase/decrease its transmission rate based on the response/feedback from the Fuzzy Congestion Controller.

In short, the Fuzzy Traffic Controller's goal is to simultaneously monitor mean rate and reject bursts while at the same time prevent and relieve congestion.